

WHAT WE CLAIM IS:

1. A method for distributing data over a widely distributed network, comprising:
receiving at a master, coupled to a plurality of slaves, a slave state of a first slave of said plurality of slaves;
determining at said master that said first slave requires further data;
determining at said master, in accordance with a predefined bandwidth resource allocation method, a second slave of said plurality of slaves from which to transfer said further data to said first slave;
scheduling at said master, a data transfer of said further data from said second slave to said first slave; and
transmitting instructions from said master to said first slave, instructing said first slave to obtain said further data from said second slave.
2. The method of claim 1, wherein said local state of said first slave is a saved current state of file storage of said first slave including an indication of files that said first slave requires and an indication of what portion, if any, of a particular one of the files that said first slave has already stored in memory.
3. The method of claim 1, wherein said determining said second slave further comprises:
determining which available slaves of said plurality of slaves have said further data;
ascertaining system resources; and
establishing that said second slave is an preferred slave, from said available slaves, from which to transfer said further data to said first slave, based on said system resources.
4. The method of claim 3, wherein said ascertaining further comprises:
determining differences between maximum system resources system and resources in use.
5. The method of claim 1, wherein said scheduling further comprises:
determining overall system resources; and
establishing that said second slave is a preferred slave from which to transfer said further data to said first slave based on said overall system resources.
6. The method of claim 1, wherein said scheduling further comprises:

determining incremental system resources that have changed since last determining said incremental system resources; and

establishing that said second slave is a preferred slave from which to transfer said further data to said first slave based on said incremental system resources.

7. The data distribution system of claim 1, wherein said scheduling further comprises:
determining partition system resources of a partition of system components that includes said first slave; and

establishing that said second slave is a preferred slave from which to transfer said further data to said first slave based on said partition system resources.

8. The method of claim 1, further comprising, at any time before said scheduling:
determining at said master, overall system resources;
computing at said master, a preferred data transfer schedule based on said overall system resources;

storing said data transfer schedule on said master, such that said scheduling further comprises selecting said instructions from said data schedule.

9. The method of claim 1, further comprising, before said receiving, obtaining a file set at said first slave, where said file set is a list of files and version numbers of said files that said first slave must obtain.

10. The method of claim 9, further comprising, after said obtaining:
ascertaining at said first slave, said state of said first slave; and
transmitting said state to said master.

11. The method of claim 10, wherein said state is ascertained by establishing what further data said first slave still requires.

12. The method of claim 11, wherein said establishing further comprises subtracting data already stored on said first slave from that required per said file set.

13. The method of claim 1, further comprising:
receiving said instructions at said first slave;
requesting said further data from said second slave;
receiving at said first slave said further data from said second slave; and
storing said further data on said first slave.

14. The method of claim 13, further comprising:
ascertaining at said first slave, a revised state of said first slave; and
transmitting said revised state to said master.
15. The method of claim 14, wherein said ascertaining further comprises determining whether said transfer of said further data from said second slave to said first slave was successful.
16. The method of claim 15, wherein if said transfer of said further data was unsuccessful:
assessing at said first slave failure information identifying components of said WAN responsible for said unsuccessful transfer; and
storing said failure information in said state, such that said failure information is transferred to said master.
17. The method of claim 16, further comprising:
incrementing a count for each component of said WAN that has been identified as being responsible for an unsuccessful transfer;
identifying a component of said WAN as failed when said count reaches a predetermined number.
18. The method of claim 13, further comprising verifying said further data.
19. The method of claim 13, further comprising:
determining that no more data is required for a particular file; and
updating said state to show that no more data is required for said particular file.
20. The method of claim 1, further comprising, before said receiving:
generating said further data at a generator;
generating a checksum of said further data at said generator; and
sending said further data and said checksum to said second slave.
21. The method of claim 20, further comprising, before said sending:
generating a final flag at said generator, where said final flag indicates that said further data is a final block of data for a particular file, such that said final flag is sent to said second slave together with said further data and said checksum.
22. The method of claim 21, further comprising:

receiving said further data at said second slave; and
storing said further data on said second slave.

23. The method of claim 22, further comprising, after said second slave receives said further data, verifying said further data based on said checksum.

24. A system for distributing data, comprising:
a plurality of dispersed slaves configured to store and serve data;
at least one generator configured to generate said data and to make said data available to a subset of said plurality of slaves; and

a master coupled to said plurality of dispersed slaves, where said master is configured to:

state information from a first slave of said plurality of slaves;

determining that said first slave requires further data;

determining at said master, in accordance with a predefined bandwidth resource allocation method, a second slave of said plurality of slaves from which to transfer said further data to said first slave;

schedule a data transfer of said further data from said second slave to said first slave;

and

transmit instructions from said master to said first slave, instructing said first slave to obtain said further data from said second slave.

25. The system of claim 24, wherein said master comprises:

a processor;

communications circuitry; and

a memory, comprising:

a operating system;

communication procedures for communicating with said plurality of slaves via said communication circuitry;

an indication maximum system resources;

an indication of system resources in use;

an indication of a state of each slave; and

a scheduler for scheduling data transfer between slaves.

26. The system of claim 25, wherein each slave of said slaves further comprises:

a processor;
communications circuitry; and
a memory, comprising:
a operating system;
communication procedures for communicating with said plurality of slaves and said master via said communication circuitry;
a file set file containing a list of files required by said slave; and
at least a portion of one of said files and an indication of how much of said file is still required.

27. The system of claim 25, wherein said master and each of said slaves preferably include verification procedures for verifying said data.

28. A method for distributing a plurality of files over a network, comprising:
at each slave of a plurality of slaves, determining a current file length for each of the plurality of files, the current file length of each respective file representing an amount of the respective file currently stored at the slave, and sending slave status information to a master, the slave status information including the current file length for each of the plurality of files;
at the master, receiving the slave status information from each of the slaves;
at the master, scheduling copy operations based on the slave status information received from the slaves, each scheduled copy operation for copying a specified portion of a specified file of the plurality of files, from a first specified slave of the plurality of slaves to a second specified slave of the plurality of slaves; and
at the master, storing bandwidth capability information indicating data transmission bandwidth capabilities for a plurality of resources required for transmitting data between the plurality of slaves, and storing bandwidth usage information indicating amounts of data transmission bandwidth allocated to copy operations scheduled by the master, the stored bandwidth usage information indicating a total allocated bandwidth for each resource of the plurality of resources;
the scheduling including allocating for each copy operation an amount of data transmission bandwidth with respect to each resource required by the copy operation, and updating the stored bandwidth usage information in accordance therewith;
the scheduling further including only scheduling copy operations that do not cause the total allocated bandwidth of any resource to exceed the bandwidth capability of that resource.

29. The method of claim 28, wherein a first file of the plurality of files includes a first set of blocks and a second set of blocks, the method including:

executing the copy operations scheduled by the master; and

at one of the slaves, generating the second set of blocks of the first file after at least one block of the first file has been copied to another one of the slaves by the execution of a respective copy operation scheduled by the master.

30. The method of claim 28, wherein the scheduling includes applying a cost function to a plurality of potential copy operations to produce a respective score for each potential copy operation, selecting a copy operation of the plurality of potential copy operations having a best score, and scheduling the selected copy operation.